The effect of facial feedback on the evaluation of statements describing everyday situations and the role of awareness

Article (Accepted Version)

Running head: Facial Feedback and Evaluations

The effect of facial feedback on the evaluation of statements describing everyday situations and the role of awareness

Jakob Kaiser [a] & Graham C.L. Davey [b]
[a] Corresponding Author: School of Psychology, University of Sussex, Falmer, Brighton, BN1 9QH, UK; J.Kaiser@sussex.ac.uk
[b] School of Psychology, University of Sussex, Falmer, Brighton, BN1 9QH, UK; grahamda@sussex.ac.uk

Abstract
According to theories of embodiment enacting a smile or a frown can positively or negatively influence one's evaluations, even without awareness of one's facial activity. While some previous studies found evidence for facial feedback effects, recent replication attempts could not confirm these findings. Are our decisions throughout the day amenable to the state of our facial muscles? We tested the effect of smiling and frowning on the evaluation of emotional sentences describing everyday situations. While most previous studies based their assessment of awareness on verbal debriefing interviews without explicitly defined criteria, we employed a written debriefing questionnaire in order to avoid potential bias when identifying participants' awareness. Our results indicate that smiling/frowning increased/decreased sentence ratings only for participants aware of their expressions. This emphasizes the importance of more rigorous awareness tests in facial feedback studies. Our results support the view that facial feedback cannot necessarily influence us without conscious mediation.

1 Introduction

Human beings constantly evaluate their surroundings to identify both beneficial opportunities and situations that might pose a potential threat. Theories of embodiment suggest that our evaluations are partly shaped by the states of our body (Barrett & Lindquist, 2008; Laird & Lacasse, 2013; Niedenthal, 2007). For example it has been suggested that afferent feedback from the tension in our emotion-relevant facial muscles can influence our processing of affective stimuli (McIntosh, 1996; Adelmann & Zajonc, 1989). Simply put, we might perceive our surroundings as more positive when smiling and more negative when frowning. This so-called facial feedback theory was supported by some previous studies which found that people who are enacting smiles (as opposed to either frowning or being prevented from smiling) gave more positive ratings for stimuli like cartoons (Laird, 1974; Strack et al., 1988; Dzokoto, Wallace, Peters, & Bentsi-Enchill, 2014) facial expressions (Dimberg & Söderkvist, 2011) or short video clips (Davis, Senghas, & Ochsner, 2009; Soussignan, 2002). However, a recent replication project of several laboratory groups found no effect of facial expressions on the evaluation of cartoons, despite a sample size of altogether about 1.000 participants (Wagenmakers, Beek, Dijkhoff, & Gronau, 2016). Thus, there is currently a debate about the generalizability of earlier found facial feedback effects. Are our decisions throughout the day really amenable to the state of our facial muscles? The current experiment tests the effect of facial muscle activation (smiling and frowning) on the evaluation of sentences describing everyday situations. These situations were chosen to be either clearly positive/negative or ambiguous in their affective value, in order to see if facial influence is stronger in cases where the valence value of a situation is less determined. Since the facial feedback hypothesis proposes that the effect of facial activation does not depend on the conscious appreciation of one's facial expressions, we compared the extent of facial feedback effects between participants with and without awareness of their emotional expressions.

Previous studies employed the evaluation of short statements describing common life events as an approximation of everyday decision making, especially with respect to trait-related interpretation biases. It was found that people with high levels of trait anxiety (Hirsch & Mathews, 1997; Voncken, Bögels, & De Vries, 2003; Wenzel & Lystad, 2005) or diagnosed anxiety disorders (Eysenck, Mogg, May, Richards, & Mathews, 1991; Stopa & Clark, 2000) tend to interpret especially ambiguous statements as more negative. Some findings even suggest that such interpretation biases might help to maintain emotional disorders (Hayes, Hirsch, Krebs, & Mathews, 2010; Murphy, Hirsch, Mathews, Smith, & Clark, 2007; Salemink, van den Hout, & Kindt, 2010; Stopa & Clark, 2000). While such maladaptive interpretative tendencies are often explained with respect to cognitive mechanisms (e.g. Behar, DiMarco, Hekler, Mohlman, & Staples, 2009; Clark, 2001; Mathews & MacLeod, 2005), embodiment theories propose that body activity is important for affective processing (Barrett & Lindquist, 2008; Winkielman, Niedenthal, Wielgosz, Eelen, & Kavanagh, 2015). The current study tests the possibility that manipulating participants' facial muscle activation might influence the rating of verbal descriptions of events similar to those employed in experiments measuring pathologically-relevant interpretation biases. Since some theories of embodiment propose that bodily feedback might be especially relevant when making decisions under uncertainty (Bechara & Damasio, 2005), we test the effect of facial feedback on both clearly valenced and ambiguous descriptions. As the latter are less determined in their emotional value, one could suspect that any influence due to changes in facial expressions are stronger in this case.

Investigating the role of facial feedback for the processing of affective sentences is also relevant for theories of languages processing. Embodiment accounts suggests that understanding a sentence might partially rely on bodily feedback related to its meaning (Barsalou, 1999; Buccino et al., 2005; Zwaan, 2004). In line with this view it has been found that selective suppression of facial feedback during the reading of emotional sentences (e.g.

blocking of smiles for positive statements) can lead to an increase in comprehension time (Havas, Glenberg, Gutowski, Lucarelli, & Davidson, 2010; Havas, Glenberg, & Rinck, 2007), as well as EEG responses indicating increased processing effort (Davis, Winkielman, & Coulson, 2015). While these studies mainly measured the effect of facial activity on time and effort needed to comprehend emotional sentences, it is less clear if muscle tension affects the actual interpretative outcome as well. Davis et al. (2015) for example found no effect of facial feedback on subsequent ratings of target sentences. Since their procedure was optimized for EEG measurements during reading, the ratings were performed only after a delay of several seconds when emotion-relevant facial activation was not enacted anymore which might have diminished the effect on the ratings. Thus, the current study more directly tests the role of facial feedback on evaluative outcomes by ensuring that facial activation is present during both the reading and rating of emotional sentences.

One important goal of the current study was to investigate the role of task-awareness for the facial feedback effect on evaluations. It has often been considered crucial for the interpretation of feedback effects if participants are aware that they currently perform emotional expressions (McIntosh, 1996; Zajonc, Murphy, & Inglehart, 1989). If facial actions could influence participants without them being aware of their expressions, then this would suggest that facial muscle tension can directly shape one's interpretative tendencies via bodily afferent feedback. If such effects were only found for participants aware of their expressions, then such results might be explained either due to demand or expectancy effects, or due to the activation of expression-congruent associations in the participant (e.g., the instruction to perform a smile could function as a prime of the concept of happiness; Buck, 1980; Laird, 1974; Zajonc et al., 1989). Thus, one challenge of facial feedback studies has been to demonstrate feedback effects in the absence of conscious knowledge about the emotional relevance of one's current expressions. Note that this would still allow for participants to have conscious changes in their emotional states (e.g. they might be happier during smiling), as long as they are not aware that these changes are supposed to be elicited by their facial actions. Previous studies used non-emotional cover-stories in order to avoid participants becoming aware of the emotional relevance of their expressions. Facial feedback effects were found after the exclusion of task-aware participants in most previously published reports (e.g. Dimberg & Söderkvist, 2011; Ito, Chiao, Devine, Lorig, & Cacioppo, 2006; Strack et al., 1988), but not all of them (Reisenzein & Studtmann, 2007; Wagenmakers et al., 2016). Given the relevance of awareness for understanding such findings, it is important to consider how task awareness should be assessed. Almost all facial feedback studies report using a verbal debriefing interview in order to measure participants' task awareness. However, research on experiment compliance suggests that participants are less likely to admit experiment-relevant knowledge in verbal as compared to questionnaire-based debriefings (Newberry, 1973). Most publications concerning facial feedback do not report a standardized, pre-defined set of debriefing questions or state explicit criteria for classifying the answers as showing task awareness. This make it hard for other researchers to critically evaluate the claims of researchers concerning the awareness of their participants. Importantly, debriefings interviews are themselves susceptible to demand effect (Orne, 1962), both on the side of the experimenter (who might be implicitly biased in wanting to succeed with the cover story), and the participant (who can be reluctant to contradict the 'authoritative' cover story in a direct conversation; Blackhart, Brown, Clark, Pierce, & Shell, 2012). As a possible improvement in this respect, Davis et al., (2009) and Soussignan (2004) used a written debriefing questionnaire and explicit criteria for identifying task-aware participants. While these studies excluded task-aware participants from their analysis, the current study employs this approach for a direct comparison of aware and unware participants. This will clarify how far any effect of facial feedback on verbal statements might depend on participants' conscious knowledge about their facial expressions. As the approach for identifying aware participants can vary considerably between studies, probing the role of task awareness might be an important step in explaining the divergent in study results concerning facial feedback.

Overall, the current study assessed the effect of smiling, frowning or (as a control) no specific facial movements on the affective evaluations of a range of statements involving everyday situations that were either positive, negative or ambiguous in their emotional valence. The main dependent variable of interest where participants' ratings of these situations while performing the intended facial expressions. In order to ensure that the emotional muscles were selectively activated during the ratings, we also measured facial muscle activation of the corrugator supercilii ('frowning') and the zygomaticus major ('smiling') muscle via electromyography (EMG; Tassinary, Cacioppo, & Vanman, 2000). We motivated the facial movements with a non-emotional cover story, but a structured debriefing questionnaire was employed that would allow comparison of effects in aware and unaware participants participants. The facial feedback hypothesis predicts that smiling/frowning leads to a more positive/negative evaluation of the situations irrespective of whether they are aware of that feedback. As the meaning of ambiguous situations should be more amenable to contextual influences, the effect could be expected to be more pronounced here. The comparison of task-aware und task-unaware participants allows to determine if any effect of facial feedback on sentence evaluations is dependent on a conscious mediation such as expectancy effects or activation of emotion-congruent associations.

2 Method

2.1 Participants

Participants were 121 undergraduate students (93 female) from Sussex University, taking part for course credit or financial reimbursement of £5. Mean age was 19.97 (SD = 2.88).

Materials

An initial set of 50 sentences describing either positive, negative, or ambiguous situations were created. In a pilot study, 32 participants rated the valence of each sentence on a 9-point-scale ranging from -4 (*negative*) to +4 (*positive*). Based on these ratings, we chose 9 clearly positive situations (each with a mean rating greater than 2, overall M = 3.66, SD = 0.13; e.g. "I've had a really productive day and am looking forward to a relaxing evening."), 9 clearly negative sentences (each with a mean rating less than -2, M = -3.69, SD = 0.44; e.g. "I have so much work already and now I hear we have another test next week."), and 9 ambiguous sentences (each having a mean rating between -1 and +1, M = 0.20, SD = 0.44; e.g. "The person from the other table looked over to me again and again").

2.2 Procedure

Participants were informed that the purpose of the study was to investigate how the activation of different muscle groups around the eyes might influence reading speed (cf. Topolinski & Strack, 2009, Experiment 4). For this reason they would be expected to perform reading tasks while contracting either muscle groups directly above or under the eye. Initially, they were instructed to lower the eyebrows whenever the message *above eye* would be shown on the screen, and raise the cheeks upwards and outwards as much as possible following the message *below eye*. These instructions had been identified in piloting trials as leading to typical frowning and smiling movements respectively. Note that in accordance with previous studies (e.g. Dimberg & Söderkvist, 2011) *frowning* in our experiment refers to a movement of the brows only. Our instructions for smiling did not include any direct mention of the lips, as referring to a body part directly under the eyes was more consistent with our cover story. In a short training phase participants practiced the movements while fixating a cross in the middle of the screen (presented to them as a baseline measurement). The experimenter corrected participants' movements when necessary until participants produced the desired expressions. Participants were told to hold their muscle movements until they received new instructions on the screen, and not to make any of the two movements when the message *neither* was shown.

The main phase of the experiment consisted of three blocks. In each block all of the 27 situations were presented once, with an equal number of each type assigned to one of the three movement instructions. The initial assignment of movements to situation was randomized and counterbalanced between the blocks, i.e. each sentence was rated once under each of the three facial movement conditions. Each trial started with the display of a movement instruction (*above eye*, *below eye* or *neither*) for one second, followed by the presentation of one of the situations. After three seconds a straight line with the two anchors *negative* and *positive* appeared below the sentences, functioning as a continuous rating scale from -420 to +420. Participants used the mouse to select a position on this scale reflecting their evaluation of the situation described. They were told that these ratings were necessary in order to ensure that they truly processed the meaning of each sentence. In order to make sure that participants would not memorize specific evaluation decisions, they did not receive any numerical feedback about their rating. After each trial the instruction *neither* appeared and a three-second pause followed to allow for the relaxation of the facial muscles. Thus, participants performed the facial movements during both the reading and the rating of each situation. Additionally, the experiment contained 15 distractor trials per block, during which instead of a sentence a random letter combination and a moving dot (ostensibly representing a typical reading pattern) were displayed. Participants were instructed to follow this dot with their eyes. While they were told that this was meant to identify effects of muscle activation during non-semantic processing, these trials only functioned to reinforce the cover story, and had no relevance for the main task.

In order to evaluate participants' awareness that their facial muscle activation was related to emotional expressions, two approaches were tested: An open-response task, were participants had to formulate the study purpose in their own words, and a prompted awareness task which asked participants to select applicable study purposes from a list of possibilities. More specifically, a two-page debriefing questionnaire was administered after the main task. On the first page, participants were instructed to write down what they believed to have been the purpose of the experiment. No further instructions were given and no upper or lower limit for the length of the answer was indicated. The second page of the questionnaire included a prompted awareness task including descriptions of 13 items that may or may not be relevant to experiment. Some of these items related to the cover story of the current study (e.g. "the pattern of your eye movements"), some described aspects neither relevant to the cover story or the actual study purpose (e.g. "your gender"), and, importantly two target items were directly relevant to the true goal of the experiment ("the emotions related to your facial expressions" and "your mood"). Participants were asked to mark all the items on the list that they believed to be relevant to the experiment. An indication of awareness during an open response might be seen as a stricter criterion of awareness, since here the participant has to formulate the study purpose on her own. Choosing the actual purpose from a given list can be seen as a more liberal criterion (since reading the actual purpose after the experiment might actually trigger awareness), but could potentially have the benefit of identifying aware participants that were unwilling or unable to formulate the study purpose on their own.

2.3 EMG Measurements

Two pairs of 4mm Ag-AgCl electrodes were used for recording EMG activity from the corrugator and zygomaticus muscles. Electrode placement followed the recommendations of Fridlund & Cacioppo (1986). The signal was recorded by a Biopac Mp36 measurement unit at

a sample rate of 2000 Hz. A low pass filter of 500 Hz was employed to attenuate noise artefacts. The signal was further treated to a moving-average integrator with a window size of 100 samples.
For each trial the mean EMG activation of the three second display of the sentence was calculated and the mean value of one second directly before trial onset (i.e. before the movement instruction appeared) was subtracted from this value. The resulting values were transformed to z-scores.

3 Results

3.1 Assessment of Task-Awareness

First the free response section of the debriefing questionnaire was evaluated. Participants were classified as task-aware if any mention was made of a) facial movements in the task being related to emotional expressions, b) the expressions being intended to or actually influencing affective states, or c) the expressions being intended to or actually influencing the sentence ratings. Following these criteria fifty-four participants were considered to be task-aware. Secondly, participants' responses to the list of possible study purposes was assessed. Here they were classified as showing 'prompted awareness' if they indicated that either of the true purposes of the study were relevant. Twenty-six participants showed only prompted awareness, i.e. they did not indicate any signs of awareness in their text but when subsequently presented with a list of possible study purposes checked at least one of the relevant target items. Forty-one participants neither indicated task-awareness in their written statements nor on the subsequent checkbox list and hence where considered clearly task-unaware.

3.2 Manipulation Check

EMG data was used to test the success of the movement instructions in generating appropriate facial muscle activation (s. Figure 1). We also evaluated the possibility that a difference in task awareness might have led to different degrees of muscle movements. For both the corrugator and the zygomaticus, we conducted separate 2-way mixed ANOVAs with the within factor instruction (*above eye*/*below eye*/*neither*) and the between factor awareness level (no awareness/prompted awareness/full awareness). Here and in the subsequent analyses, Bonferroni corrections were used for post-hoc tests. For both muscles we observed a main effect of instruction, zygomaticus: $F(2, 236) = 1107.43$, $p < .001$, $\eta_p^2 = .90$, corrugator: $F(2, 236) = 2969.50$, $p < .001$, $\eta_p^2 = .96$. In both cases there was no main effect of awareness level, zygomaticus: $F(1, 118) = 0.16$, $p = .85$, $\eta_p^2 < .01$, corrugator: $F(1, 118) = 1.84$, $p = .16$, $\eta_p^2 = .03$, nor an instruction*awareness interaction, zygomaticus: $F(4, 236) = 0.71$, $p = .59$, $\eta_p^2 = .01$, corrugator: $F(2, 236) = 0.37$, $p = .82$, $\eta_p^2 = .01$. This indicates that the instructions lead to different degrees in smiling and frowning with no difference in their effect between unaware and aware participants. The *below eye* instruction lead to significantly more zygomaticus activity than the neither instruction, $t(120) = 38.00$, $p < .001$, $d = 3.45$, and the *above eye* instruction, , $t(120)= 33.91$, $p < .001$, $d = 3.08$. There was also less zygomaticus activity in the *neither* than in the *above eye* condition, $t(120)= -5.13$, $p < .001$, $d = 0.47$. This might be an indication that the neither instruction lead to an overall relaxation of the face. Importantly, the change of in zygomaticus activity to from baseline was negative in both above eye and neither condition, and only positive after the below eye instruction (cf. Figure 1). For the corrugator, the *above eye* instruction lead to significantly more activity than the *neither* instruction, $t(120)= 66.10$, $p < .001$, $d = 6.01$, and the *below eye* instruction, $t(120)= 68.98$, $p < .001$, $d = 6.27$, which did not significantly differ from each other, $t(120)= -1.06$, $p > .99$, $d = -0.10$. Thus, the movement instructions selectively increased smiling in the *below eye* condition and frowning in the *above eye* condition.

3.3 Effects of Facial Expressions on Sentence Ratings

Sentence ratings were z-transformed separately for each participant (see Table 1). A 3-way expression (smiling/frowning/neither) x situation type (negative/positive/ambiguous) x awareness level (unaware/prompted awareness/clearly aware) mixed ANOVA was conducted on the ratings. This revealed a main effect of situation, $F(2, 236) = 3474.87$, $p < .001$, $\eta_p^2 = .97$, expression, $F(2, 236) = 6.75$, $p = .01$, $\eta_p^2 = .05$, but no main effect of awareness, $F(2, 118) < .001$, $p > .90$, $\eta_p^2 < .001$. There was a situation*expression interaction, $F(4, 472) = 3.20$, $p = .01$, $\eta_p^2 = .03$, an awareness*expression interaction, $F(4, 118) = 9.42$, $p < .001$, $\eta_p^2 = .14$, but no situation*awareness interaction, $F(4, 118) = 0.64$, $p = .63$, $\eta_p^2 = .01$. All of these effects were qualified by a situation*expression*awareness interaction, $F(8, 472) = 3.15$, $p = .002$, $\eta_p^2 = .05$. These results indicate a difference in the effects between participants of different awareness levels. Thus, for each awareness level a 2-way expression (smiling/frowning/neither) x situation type (negative/positive/ambiguous) within ANOVA was conducted. There was a main effect of situation for all groups, unaware: $F(2, 80) = 786.16$, $p < .001$, $\eta_p^2 = .95$, prompted awareness: $F(2, 50) = 952.18$, $p < .001$, $\eta_p^2 = .97$, clearly aware: $F(2, 106) = 2893.67$, $p < .001$, $\eta_p^2 = .98$. For both the clearly unaware and prompted awareness group there was neither a main effect of expression, clearly unaware: $F(2, 80) = 0.37$, $p = .69$, $\eta_p^2 = .01$, prompted: $F(2, 50) = 1.22$, $p = .31$, $\eta_p^2 = .05$, nor an expression*situation interaction, clearly unaware: $F(4, 160) = 1.23$, $p = .30$, $\eta_p^2 = .03$, prompted: $F(2, 100) = 0.72$, $p = .58$, $\eta_p^2 = .03$. For clearly aware participants, there was a main effect of expression, $F(2, 106) = 17.30$, $p < .001$, $\eta_p^2 = .25$, as well as an expression*situation interaction, $F(4, 212) = 8.89$, $p < .001$, $\eta_p^2 = .14$. Thus, in all awareness classes the ratings depended on the type of situation, but only for clearly aware participants the rating depended on the expression performed during the evaluation. In all classes, ratings were significantly higher for positive situations than ambiguous and negative ones, and lower for negative than ambiguous situations (all p's <.01, s. Table 1 for means and standard deviations). For the clearly aware group only, smiling lead to higher ratings than no expression $t(53)= 3.73$, $p < .001$, $d = 0.51$ and frowning, $t(53)= 4.36$, $p < .001$, $d = 0.59$. Frowning lead to lower ratings than no expression, $t(53)= -3.93$, $p < .001$, $d = 0.53$. The expression*situation interaction suggested that for some situation types the effect of the expression might have been stronger. In order to be able to compare the degree of influence in ratings produced through different expressions in the aware participants, we calculated for each situation type a change score by subtracting rating during frowning from the ones during smiling. The higher this score, the more amenable to facial feedback a sentence type could be considered. Difference scores for ambiguous situations (m = 0.23, SD = .38) were significantly higher than for positive situations (m = 0.08, SD = .22), $t(53) = 3.95$, $p < .001$, $d = 0.54$, or negative ones (m = 0.11, SD = .19), $t(53) = 3.41$, $p = .004$, $d = 0.46$), with no difference between the latter two, $t(53)= -1.40$, $p = .17$, $d = -0.19$. Thus, while all sentences were influenced by the facial expressions, this effect seemed to have been most pronounced for descriptions of ambiguous situations. Overall, while in all groups ratings depended on the valence of the situation, facial feedback only had an influence for clearly aware participants. The influence here was strongest for ambiguous situations. Given that unaware participants did not show a facial feedback effect, it would be important to know if this result might be due to lack of statistical power (i.e. the number of unaware participants being insufficient) or if it can be seen as positive evidence for the absence of an expression effect. We explored this question with a Bayes factor analysis (Dienes, 2014; Jarosz & Wiley, 2014; Wagenmakers, 2007). A Bayes factor (BF) is a numerical value that can differentiate between positive evidence for a hypothesis (usually BF > 3, s. Jarosz & Wiley, 2014 for recommended interpretation thresholds), positive evidence against it (BF < 1/3) and not entirely conclusive data (due to a lack of statistical power, 1/3 < BF < 3). Since the prompted awareness group showed the same pattern of results as the clearly unaware group, suggesting that participants in this group were actually unaware during the experiment, both groups were considered to be unaware and hence included in this analysis. Using JASP (JASP Team, 2016), we calculated the Bayes factor testing the main effect of expression on the sentence ratings, as well as the Bayes factor for the inclusion of the expression*situation interaction (as compared to a model that only included the main effects). For both expression, BF = 0.171, and the expression*situation interaction, BF = 0.008, the resulting Bayes factors were relatively low. This suggest that our data provides evidence against an effect of expressions on the ratings. Thus, the absence of an expression influence in unaware participants is not likely to be due to a lack of statistical sensitivity, but provides positive evidence for the null hypothesis, i.e. that muscle activation failed to influence the situation evaluations.

4 Discussion

The current study tested the effect of emotional facial muscle activation (smiling and frowning) on the evaluation of sentences describing clearly positive/negative or ambiguous situations. We employed a written debriefing questionnaire with explicit criteria for probing how far any affect might depend on participants' awareness of the emotional relevance of their facial movements. It was found that for participants with awareness, facial activation did influence the evaluation of the described situations, with ambiguous situations being influenced significantly stronger by facial activity. There was no effect for participants without awareness of their expressions. As there was no difference in muscle activation between unaware and aware participants, these results cannot be explained with a difference in task compliance between the groups. Additionally, Bayesian analysis of the data indicates that the number of unaware participants was large enough to interpret the results as positive evidence for a null effect, i.e. a higher number of unaware participants is not likely to have produced a different finding.

Methodologically, these results demonstrate the importance of finding a reliable way to identify task-awareness in research concerning facial feedback. Only participants classified as aware based on their free-text answers showed an effect of facial actions on ratings. Participants who only indicated awareness when directly presented with the true study purpose (i.e. prompted awareness) did not differ in their effect from clearly unaware participants, suggesting that this approach for identifying awareness can lead to an overestimation of aware participants because the answers here might represent participants' post-experimental beliefs elicited by the mentioning of the actual purpose. One limitation of the current study is that the debriefing questionnaire does not explain why awareness of the emotional relevance of the expressions leads to a systematic change in ratings. One possibility would be that participants might have reacted to perceived experimental demand. It would also be conceivable that consciously considering one's expression would lead to activation of conceptual assumptions about the relation between certain expressions and emotions (Barrett, 2006; Bem, 1972; Schnall & Laird, 2003). Thus, the conscious identification of one's expressions might function as a contextual prime, leading to a biased interpretation of the target situations. Future studies could try to distinguish between different possible cognitive mechanism that might mediate the relation between task-awareness and the behavioural outcome, for example by assessing participants' conceptual beliefs about the influence of expressions on their emotional states.

Overall, several studies exist that report facial feedback effects in as unaware identified participants, using a range of different experimental designs and outcome measures (e.g. Andréasson & Dimberg, 2008; Dimberg & Söderkvist, 2011; Ito et al., 2006; Strack et al., 1988). This support the idea that facial muscle tension can causally influence our affective processing at least in some aspects even in the absence of awareness. However, a recent replication attempt of Strack et al. (1988) could not find an effect of facial actions on evaluations, despite a high overall sample size of about 1700 participants (Wagenmakers et al., 2016). This suggests that further research is needed to establish the reliability of facial feedback effects. The current study extends these previous findings by demonstrating the considerable influence of task awareness on effects of facial manipulations. It is noteworthy that many feedback studies use verbal debriefing interviews with no explicit criteria for their classification (or sometimes report no awareness test for their cover story at all). Thus, there might be a risk of a bias for classifying aware as unaware participants in some studies. Our results suggests that such a bias, combined with the currently widely discussed tendency not to publish null results (Ferguson & Heene, 2012; Kühberger, Fritz, & Scherndl, 2014) might lead to an overestimation of the extent and intensity of facial feedback effects.

Note that our results do not imply a general rejection of the possibility of unconscious facial feedback effects. While previous studies finding facial feedback effects without awareness employed mainly pictorial stimuli (e.g. Davis et al., 2009; Dimberg & Söderkvist, 2011; Soussignan, 2004), we found no effect on ratings of verbal statements describing relatively common life events, corroborating a similar finding by Davis et al. (2015). Other studies investigating facial effects on emotional sentences did not directly measure evaluative outcomes, but found that selectively blocking smiling/frowning increased reading times of emotion-congruent sentences (Havas et al., 2010, 2007). The divergent results concerning the presence and absence of such effects might suggest that proponents of the facial feedback hypothesis would have to identify in more specific terms what kind of evaluative processes can be influenced by facial activation. For example, dual-process models of decision making often distinguish between intuitive (usually relatively quick) and analytical (usually more elaborated) decisions (Chen & Chaiken, 1999; Evans & Stanovich, 2013). Some theories of embodiment suggest that bodily feedback might function as an additional heuristic for decisions where external information is not readily available, effectively allowing for intuitive 'gut feeling' decisions (Damasio, 1996; Sütterlin, Schulz, Stumpf, Pauli, & Vögele, 2013). On the other hand, language processing has sometimes been linked to more analytical processing (cf. Evans, 2008; Sadler-Smith, 2011). It might be possible that the task to evaluate sentences in the current study lead to a more elaborated or analytical decision making approach in the participants than the evaluation of pictorial stimuli and thus made it more likely to base the ratings on cognitive strategies rather than facial feedback. The effect on reading times in earlier studies could be interpreted a trade-off between processing effort and amenability to facial feedback, e.g. maybe participants invest processing resources into tuning out potentially misleading bodily feedback in these cases. Future studies could address this distinction by contrasting the effect of facial muscle activation under conditions that encourage more or less elaborated decision making, e.g. by manipulating the time available for forming a judgment. It is also noteworthy, that some studies suggested that facial feedback effects might depend on interpersonal differences (Andréasson & Dimberg, 2008; Laird et al., 1994). Thus, it would be possible that people differ in their amenability towards bodily feedback. If that was the case, it would be important to specify the personality factors that determine the responsiveness towards afferent facial feedback.

Overall, the current study is in line with a limited view of facial feedback effects, suggesting that facial feedback is not necessarily a determinant of evaluative outcomes (Reisenzein & Studtmann, 2007; Wagenmakers et al., 2016), which leaves open the possibility that its influence is relevant under special circumstances (cf. Davis et al., 2015; Maringer, Krumhuber, Fischer, & Niedenthal, 2011). Experimentally distinguishing between affect-related processes that are or are not dependent on bodily feedback such as the state of our facial muscles, will be necessary in order to further our understanding of the role of embodiment in our everyday life.

5 References

Adelmann, P. K., & Zajonc, R. B. (1989). Facial efference and the experience of emotion. *Annual Review of Psychology*, *40*, 249–280. https://doi.org/10.1146/annurev.psych.40.1.249

Andréasson, P., & Dimberg, U. (2008). Emotional Empathy and Facial Feedback. *Journal of Nonverbal Behavior*, *32*(4), 215–224. https://doi.org/10.1007/s10919-008-0052-z

Barrett, L. F. (2006). Solving the emotion paradox: categorization and the experience of emotion. *Personality and Social Psychology Review : An Official Journal of the Society for Personality and Social Psychology, Inc*, *10*(1), 20–46. https://doi.org/10.1207/s15327957pspr1001_2

Barrett, L., & Lindquist, K. (2008). The embodiment of emotion. In G. R. Semin & E. R. Smith (Eds.), *Embodied Grounding: Social, Cognitive, Affective, and Neuroscientific Approaches.* Cambridge University Press. https://doi.org/http://dx.doi.org/10.1017/CBO9780511805837.011

Barsalou, L. W. (1999). Perceptual symbol systems. *The Behavioral and Brain Sciences*, *22*(4), 577-609–60. https://doi.org/10.1017/S0140525X99252144

Bechara, A., & Damasio, A. R. (2005). The somatic marker hypothesis: A neural theory of economic decision. *Games and Economic Behavior*, *52*(2), 336–372. https://doi.org/10.1016/j.geb.2004.06.010

Behar, E., DiMarco, I. D., Hekler, E. B., Mohlman, J., & Staples, A. M. (2009). Current theoretical models of generalized anxiety disorder (GAD): Conceptual review and treatment implications. *Journal of Anxiety Disorders*, *23*(8), 1011–1023. https://doi.org/10.1016/j.janxdis.2009.07.006

Bem, D. J. (1972). Self-Perception Theory. *Advances in Experimental Social Psychology*. https://doi.org/10.1016/S0065-2601(08)60024-6

Blackhart, G. C., Brown, K. E., Clark, T., Pierce, D. L., & Shell, K. (2012). Assessing the adequacy of postexperimental inquiries in deception research and the factors that promote participant honesty. *Behavior Research Methods*, *44*(1), 24–40. https://doi.org/10.3758/s13428-011-0132-6

Buccino, G., Riggio, L., Melli, G., Binkofski, F., Gallese, V., & Rizzolatti, G. (2005). Listening to action-related sentences modulates the activity of the motor system: A combined TMS and behavioral study. *Cognitive Brain Research*, *24*(3), 355–363. https://doi.org/10.1016/j.cogbrainres.2005.02.020

Buck, R. (1980). Nonverbal behavior and the theory of emotion: the facial feedback hypothesis. *Journal of Personality and Social Psychology*, *38*(5), 811–824. https://doi.org/10.1037/0022-3514.38.5.811

Chen, S., & Chaiken, S. (1999). The heuristic-systematic model in its broader context. *Dualprocess Theories in Social Psychology*. https://doi.org/10.4319/lo.2013.58.2.0489

Clark, D. M. (2001). A cognitive perspective on social phobia. In *International handbook of social anxiety: Concepts, research and interventions relating to the self and shyness* (pp. 405–430). https://doi.org/10.1016/S0006-3223(97)87445-8

Damasio, A. R. (1996). The somatic marker hypothesis and the possible functions of the prefrontal cortex. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *351*(1346), 1413–1420. https://doi.org/10.1098/rstb.1996.0125

Davis, J. I., Senghas, A., & Ochsner, K. N. (2009). How Does Facial Feedback Modulate

Emotional Experience? *Journal of Research in Personality*, *43*(5), 822–829. https://doi.org/10.1016/j.jrp.2009.06.005

Davis, J. I., Winkielman, P., & Coulson, S. (2015). Facial Action and Emotional Language: ERP Evidence that Blocking Facial Feedback Selectively Impairs Sentence Comprehension. *Journal of Cognitive Neuroscience*, *27*(11), 2269–2280. https://doi.org/doi:10.1162/jocn_a_00858

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology*, *5*(July), 1–17. https://doi.org/10.3389/fpsyg.2014.00781

Dimberg, U., & Söderkvist, S. (2011). The Voluntary Facial Action Technique: A Method to Test the Facial Feedback Hypothesis. *Journal of Nonverbal Behavior*, *35*(1), 17–33. https://doi.org/10.1007/s10919-010-0098-6

Dzokoto, V., Wallace, D. S., Peters, L., & Bentsi-Enchill, E. (2014). Attention to Emotion and Non-Western Faces: Revisiting the Facial Feedback Hypothesis. *The Journal of General Psychology*, *141*(2), 151–168. https://doi.org/10.1080/00221309.2014.884052

Evans, J. S. B. T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, *59*, 255–78. https://doi.org/10.1146/annurev.psych.59.103006.093629

Evans, J., & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science*, *8*(3), 223–241. https://doi.org/10.1177/1745691612460685

Eysenck, M. W., Mogg, K., May, J., Richards, A., & Mathews, A. (1991). Bias in interpretation of ambiguous sentences related to threat in anxiety. *Journal of Abnormal Psychology*, *100*(2), 144–50. https://doi.org/http://dx.doi.org/10.1037/0021-843x.100.2.144

Ferguson, C. J., & Heene, M. (2012). A Vast Graveyard of Undead Theories: Publication Bias and Psychological Science's Aversion to the Null. *Perspectives on Psychological Science*, *7*(6), 555–561. https://doi.org/10.1177/1745691612459059

Fridlund, A. J., & Cacioppo, J. T. (1986). Guidelines for human electromyographic research. *Psychophysiology*, *23*(5), 567–589. https://doi.org/10.1111/j.1469-8986.1986.tb00676.x

Havas, D. A., Glenberg, A. M., Gutowski, K. A., Lucarelli, M. J., & Davidson, R. J. (2010). Cosmetic Use of Botulinum Toxin-A Affects Processing of Emotional Language. *Psychological Science*, *21*(7), 895–900. https://doi.org/10.1177/0956797610374742

Havas, D. A., Glenberg, A. M., & Rinck, M. (2007). Emotion simulation during language comprehension. *Psychonomic Bulletin & Review*, *14*(3), 436–41. https://doi.org/http://dx.doi.org/10.3758/bf03194085

Hayes, S., Hirsch, C. R., Krebs, G., & Mathews, A. (2010). The effects of modifying interpretation bias on worry in generalized anxiety disorder. *Behaviour Research and Therapy*, *48*(3), 171–8. https://doi.org/10.1016/j.brat.2009.10.006

Hirsch, C., & Mathews, A. (1997). Interpretative inferences when reading about emotional events. *Behaviour Research and Therapy*, *35*(12), 1123–1132. https://doi.org/10.1016/S0005-7967(97)00069-7

Ito, T. A., Chiao, K. W., Devine, P. G., Lorig, T. S., & Cacioppo, J. T. (2006). The influence of facial feedback on race bias. *Psychological Science*, *17*(3), 256–61. https://doi.org/10.1111/j.1467-9280.2006.01694.x

Jarosz, A. F., & Wiley, J. (2014). What Are the Odds ? A Practical Guide to Computing and Reporting Bayes Factors. *Journal of Problem Solving*, *7*, 2–9.

https://doi.org/10.7771/1932-6246.1167

JASP Team (2016). JASP (Version 0.7.5.5) [Computer software]. (n.d.). Retrieved from https://jasp-stats.org/

Kühberger, A., Fritz, A., & Scherndl, T. (2014). Publication bias in psychology: a diagnosis based on the correlation between effect size and sample size. *PloS One*, *9*(9), e105825. https://doi.org/10.1371/journal.pone.0105825

Laird, J. (1974). Self-attribution of emotion: the effects of expressive behavior on the quality of emotional experience. *Journal of Personality and Social Psychology*. https://doi.org/http://dx.doi.org/10.1037/h0036125

Laird, J. D., Alibozak, T., Davainis, D., Deignan, K., Fontanella, K., Hong, J., … Pacheco, C. (1994). Individual differences in the effects of spontaneous mimicry on emotional contagion. *Motivation and Emotion*, *18*(3), 231–247. https://doi.org/10.1007/BF02254830

Laird, J. D., & Lacasse, K. (2013). Bodily Influences on Emotional Feelings: Accumulating Evidence and Extensions of William James's Theory of Emotion. *Emotion Review*, *6*(1), 27–34. https://doi.org/10.1177/1754073913494899

Maringer, M., Krumhuber, E. G., Fischer, A. H., & Niedenthal, P. M. (2011). Beyond smile dynamics: mimicry and beliefs in judgments of smiles. *Emotion*, *11*(1), 181–7. https://doi.org/10.1037/a0022596

Mathews, a, & MacLeod, C. (2005). Cognitive vulnerability to emotional disorders. *Annu Rev Clin Psychol*, *1*, 167–195. https://doi.org/10.1146/annurev.clinpsy.1.102803.143916

McIntosh, D. N. (1996). Facial feedback hypotheses: Evidence, implications, and directions. *Motivation and Emotion*, *20*(2), 121–147. https://doi.org/http://dx.doi.org/10.1007/BF02253868

Murphy, R., Hirsch, C. R., Mathews, A., Smith, K., & Clark, D. M. (2007). Facilitating a benign interpretation bias in a high socially anxious population. *Behaviour Research and Therapy*, *45*(7), 1517–29. https://doi.org/10.1016/j.brat.2007.01.007

Newberry, B. H. (1973). Truth telling in subjects with information about experiments: Who is being deceived? *Journal of Personality and Social Psychology*, *25*(3), 369–374. https://doi.org/10.1037/h0034229

Niedenthal, P. M. (2007). Embodying emotion. *Science (New York, N.Y.)*, *316*(5827), 1002–5. https://doi.org/10.1126/science.1136930

Orne, M. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist*. https://doi.org/http://dx.doi.org/10.1037/h0043424

Reisenzein, R., & Studtmann, M. (2007). On the expression and experience of surprise: no evidence for facial feedback, but evidence for a reverse self-inference effect. *Emotion (Washington, D.C.)*, *7*(3), 612–627. https://doi.org/10.1037/1528-3542.7.3.612

Sadler-Smith, E. (2011). The intuitive style: Relationships with local/global and verbal/visual styles, gender, and superstitious reasoning. *Learning and Individual Differences*, *21*(3), 263–270. https://doi.org/10.1016/j.lindif.2010.11.013

Salemink, E., van den Hout, M., & Kindt, M. (2010). How does Cognitive Bias Modification affect anxiety? Mediation analyses and experimental data. *Behavioural and Cognitive Psychotherapy*, *38*(2000), 59–66. https://doi.org/10.1017/S1352465809990543

Schnall, S., & Laird, J. D. (2003). Keep smiling: Enduring effects of facial expressions and postures on emotional experience and memory. *Cognition & Emotion*, *17*(5), 787–797. https://doi.org/10.1080/02699930244000183

Soussignan, R. (2002). Duchenne smile, emotional experience, and autonomic reactivity: A test of the facial feedback hypothesis. *Emotion*, *2*(1), 52–74. https://doi.org/10.1037/1528-3542.2.1.52

Soussignan, R. (2004). Regulatory function of facial actions in emotion processes. *Advances in Psychology Research*, *31*, 173–198.

Stopa, L., & Clark, D. M. (2000). Social phobia and interpretation of social events. *Behaviour Research and Therapy*, *38*(3), 273–83. https://doi.org/http://dx.doi.org/10.1016/s0005-7967(99)00043-1

Strack, F., Martin, L. L., & Stepper, S. (1988). Inhibiting and facilitating conditions of the human smile: a nonobtrusive test of the facial feedback hypothesis. *Journal of Personality and Social Psychology*, *54*(5), 768–77. https://doi.org/http://dx.doi.org/10.1037/0022-3514.54.5.768

Sütterlin, S., Schulz, S. M., Stumpf, T., Pauli, P., & Vögele, C. (2013). Enhanced cardiac perception is associated with increased susceptibility to framing effects. *Cognitive Science*, *37*(5), 922–35. https://doi.org/10.1111/cogs.12036

Tassinary, L. G., Cacioppo, J. T., & Vanman, E. J. (2000). The skeletomotor system: surface electromyography. In J. T. Cacioppo, L. G. Tassinary, & G. Berntson (Eds.), *Handbook of Psychophysiology* (pp. 267–299). Cambridge: Cambridge University Press. https://doi.org/http://dx.doi.org/10.1017/CBO9780511546396

Topolinski, S., & Strack, F. (2009). The architecture of intuition: Fluency and affect determine intuitive judgments of semantic and visual coherence and judgments of grammaticality in artificial grammar learning. *Journal of Experimental Psychology. General*, *138*(1), 39–63. https://doi.org/10.1037/a0014678

Voncken, M. J., Bögels, S. M., & De Vries, K. (2003). Interpretation and judgmental biases in social phobia. *Behaviour Research and Therapy*, *41*(12), 1481–1488. https://doi.org/10.1016/S0005-7967(03)00143-8

Wagenmakers, E.-J. (2007). A practical solution to the pervasive problems of p values. *Psychonomic Bulletin & Review*, *14*(5), 779–804. https://doi.org/10.3758/BF03194105

Wagenmakers, E.-J., Beek, T., Dijkhoff, L., & Gronau, Q. F. (2016). Registered Replication Report: Strack, Martin, & Stepper (1988) . *Perspectives on Psychological Science* , *11*(6), 917–928. https://doi.org/10.1177/1745691616674458

Wenzel, A., & Lystad, C. (2005). Interpretation biases in angry and anxious individuals. *Behaviour Research and Therapy*, *43*(8), 1045–54. https://doi.org/10.1016/j.brat.2004.02.009

Winkielman, P., Niedenthal, P., Wielgosz, J., Eelen, J., & Kavanagh, L. C. (2015). Embodiment of Cognition and Emotion. In *APA Handbook of Personality and Social Psychology Attitudes and Social Cognition* (Vol. 1, pp. 151–175). https://doi.org/10.1037/14341-004

Zajonc, R. B., Murphy, S. T., & Inglehart, M. (1989). Feeling and facial efference: implications of the vascular theory of emotion. *Psychological Review*, *96*(3), 395–416. https://doi.org/http://dx.doi.org/10.1037/0033-295x.96.3.395

Zwaan, R. A. (2004). The immersed experiencer: Toward an embodied theory of language comprehension. In B. H. Ross (Ed.), *Psychology of Learning and Motivation: Advances in Research and Theory* (Vol. 44, pp. 35–62). New York: Academic Press. https://doi.org/doi:10.1016/S0079-7421(03)44002-4

Table 1

Mean (and standard deviations) of the ratings for different sentence types expressed as z-scores.

	Clearly unaware (n = 41)			Prompted awareness (n = 26)			Clearly aware (n = 54)		
	Smile	None	Frown	Smile	None	Frown	Smile	None	Frown
Positive	1.119	1.141	1.125	1.143	1.152	1.154	1.163	1.101	1.085
	(0.23)	(0.21)	(0.23)	(0.14)	(0.13)	(0.12)	(0.13)	(0.14)	(0.17)
Ambiguous	-0.019	-0.037	-0.052	-0.108	-0.056	-0.077	0.093	-0.011	-0.139
	(0.21)	(0.19)	(0.20)	(0.22)	(0.22)	(0.21)	(0.23)	(0.17)	(0.25)
Negative	-1.091	-1.090	-1.087	-1.068	-1.069	-1.070	-1.049	-1.087	-1.156
	(0.18)	(0.28)	(0.23)	(0.11)	(0.14)	(0.14)	(0.17)	(0.14)	(0.12)

Figure 1

Changes in facial muscle tension (in z-scores) evoked by the movement instruction for zygomaticus (left) and corrugator (right). Error bars represent standard errors.